

O PROCESSAMENTO DE EXPRESSÕES CORREFERENCIAIS EM PORTUGUÊS BRASILEIRO: NOMES REPETIDOS, PRONOMES PLENOS E PRONOMES NULOS

PROCESSING COREFERENTIAL EXPRESSIONS IN BRAZILIAN PORTUGUESE: REPEATED NAMES, OVERT PRONOUNS, AND NULL PRONOUNS

Jefferson de Carvalho Maia
Universidade Federal de Minas Gerais

Maria Luiza Cunha Lima
Universidade Federal de Minas Gerais

RESUMO: Neste artigo, investigamos o papel de nomes repetidos, pronomes plenos e nulos no estabelecimento da coesão referencial em português brasileiro. Apresentamos resultados de um experimento de leitura autocadenciada que mostram que, nessa língua, pronomes nulos são as formas mais facilmente processadas quando da retomada de antecedentes em posição de sujeito. Comparamos esses resultados com os já obtidos para o inglês, o chinês e o espanhol e, por fim, discutimos as suas contribuições para o debate em torno do parâmetro *pro-drop* em português brasileiro.

PALAVRAS-CHAVE: Anáfora; Penalidade do nome repetido; Sujeito nulo; Parâmetro *pro-drop*.

ABSTRACT: In this paper we investigate the role of repeated names, overt and null pronouns in the promotion of referential cohesion in Brazilian Portuguese. We report the results of a self-paced reading experiment that show that, in this language, null pronouns are the most easily processed forms when referring to antecedents in subject position. We compare these results with the ones previously obtained for English, Chinese and Spanish and, in the end, we discuss their contributions to the debate on the *pro-drop* parameter in Brazilian Portuguese.

KEYWORDS: Anaphor; Repeated-name penalty; Null subject; Pro-drop parameter.

INTRODUÇÃO

Entidades introduzidas em um discurso podem ser potencialmente retomadas em enunciados subsequentes através de diferentes formas referenciais anafóricas (e.g., nomes repetidos, pronomes plenos e nulos), as quais exercem um papel fundamental no estabelecimento da coesão de um discurso. Tendo em vista que esta coesão pode ser fomentada através do uso de diferentes formas anafóricas, questões que naturalmente surgem são: o que leva os falantes a escolherem uma ou outra forma? Quais são os fatores que comandam essa escolha? E, além disso: em determinados contextos linguísticos, quais dessas formas são mais facilmente compreendidas e processadas pelos ouvintes/leitores? Este estudo constitui tentativa de resposta às essas perguntas, com a investigação, através de um experimento de leitura autocadenciada,¹ da carga de processamento de nomes repetidos, pronomes foneticamente plenos (também lexicalizados, ou simplesmente pronomes) e pronomes foneticamente nulos (também *pro*, ou simplesmente nulos), em retomadas anafóricas nas posições de sujeito e de objeto em português brasileiro (doravante PB).

Tal investigação do fenômeno da correferência, especialmente no que concerne aos sujeitos nulos, torna-se ainda mais interessante em virtude da mudança no paradigma flexional verbal pela qual o PB tem passado, o que estaria provocando o virtual desaparecimento dos

¹ Uma definição desse paradigma experimental será apresentada ao fim desta introdução e maiores esclarecimentos serão fornecidos na seção 1 deste artigo.

sujeitos nulos pessoais e constituiria uma mudança em direção a uma marcação negativa dentro do parâmetro *pro-drop*. Desse modo, os resultados desta pesquisa contribuirão tanto para a discussão sobre aspectos relacionados ao processamento mental da referência quanto para o debate em torno do parâmetro *pro-drop* em PB.

Em tentativa de resposta às questões anteriormente mencionadas, tem-se argumentado, na literatura (PRINCE, 1981; GUNDEL; HEDBERG; ZACHARSKI, 1993, *inter alia*), que os possíveis antecedentes de um termo anafórico qualquer diferem entre si em termos de proeminência, acessibilidade ou saliência cognitiva. Argumenta-se, por conseguinte, que essas diferenças de *status* cognitivo dos antecedentes têm profundo impacto na escolha das respectivas formas referenciais. Desse modo, entidades em foco tendem a ser retomadas através de formas não marcadas (e.g. pronomes plenos e nulos), ao passo que entidades que não estão em foco tendem a se realizar através de formas marcadas (e.g., descrições definidas).

Por sua vez, os fatores determinantes desses diferentes graus de acessibilidade ou saliência de referentes introduzidos em um discurso têm sido alvo de intenso debate na literatura psicolinguística. Pesquisas experimentais sobre o processamento mental da referência têm mostrado que alguns desses fatores seriam posição sintática de sujeito/objeto, posição de tópico na oração ou sentença, paralelismo sintático entre anáfora e antecedente, propriedades semânticas do verbo, prosódia (quando em situação de fala), dentre outros.

Em uma série de experimentos de leitura autocadenciada elaborados com o objetivo de investigar o comportamento de nomes repetidos e pronomes plenos no estabelecimento da coesão referencial em inglês, Gordon, Grosz e Gilliom (1993) descobriram a existência, nessa língua, de uma penalidade de processamento que relaciona forma referencial à posição sintática do antecedente, penalidade esta que foi nomeada de *repeated-name penalty* (RNP, ou penalidade do nome repetido). Em termos simples, em inglês, nomes repetidos demoram mais para ser lidos (isto é, são mais difíceis de ser processados) do que pronomes plenos quando em retomada de antecedentes em posição de sujeito, desaparecendo essa diferença de comportamento entre as duas formas anafóricas quando o antecedente se realiza em posição de objeto.²

De modo a testar a generalidade desses achados, Yang et al. (1999) investigaram, através de adaptações dos experimentos elaborados e conduzidos por Gordon, Grosz e Gilliom (1993), o processamento correferencial de nomes repetidos, de pronomes plenos e também de pronomes nulos em chinês, sistema linguístico que difere em muitos aspectos do inglês, principalmente nos de natureza sintática, entre estes a marcação do parâmetro *pro-drop*.

A definição desse parâmetro engloba uma série de propriedades, sendo as mais importantes a possibilidade de inversão do sujeito e o licenciamento de sujeitos nulos. Assim, em línguas totalmente *pro-drop*, como o chinês, é obrigatória a omissão do sujeito em determinados contextos, a rigor em sentenças de tipo afirmativa, no indicativo e com verbo finito. Já em línguas não *pro-drop*, como o inglês, o sujeito precisa estar sempre lexicalmente presente nesse tipo de sentença para que a ela seja conferido o estatuto de gramatical.³

Apesar de tais diferenças de natureza sintática entre inglês e chinês, os resultados dos experimentos conduzidos por Yang et al. (1999) revelaram um comportamento semelhante entre os dois sistemas linguísticos: o efeito de RNP também foi encontrado em chinês e, além disso, não houve diferenças estatisticamente significativas entre os tempos em que pronomes plenos e

² Esses resultados se contrapuseram a Gernsbacher (1989), para quem quanto mais explícita fosse a forma anafórica, mais rápido e efetivo seria, conseqüentemente, o estabelecimento da correferência. Assim, para a autora, a mera repetição do nome seria, em qualquer contexto, a forma mais fácil de ser processada, já que contém maior carga de informações semânticas, capazes de ativar representações mentais previamente geradas, e refere-se sem ambigüidade ao seu antecedente, sem introduzir informação nova nem exigir processos inferenciais de interpretação. Todavia, as evidências experimentais obtidas por Gernsbacher para validar suas hipóteses foram fortemente criticadas devido à metodologia por ela empregada.

³ A esse respeito, é importante levar em consideração a existência de contextos universais (i.e., independentes do parâmetro *pro-drop*) de ocorrência do sujeito nulo, como, por exemplo, o de coordenações com sujeitos correferentes.

nulos foram processados, fato que levou os autores a afirmar que eles constituem um caso de variação livre em chinês, sendo a escolha entre uma ou outra forma de natureza estilística.

Tomados em conjunto, os resultados experimentais de Gordon, Grosz e Gilliom (1993) e Yang et al. (1999) sugerem que (i) referentes introduzidos em posição de sujeito adquirem alto nível de saliência, ficando bastante acessíveis na memória discursiva, o que faz com que, nesse contexto, a carga de processamento de formas anafóricas não marcadas (pronomes) tenda a ser menor do que a de formas anafóricas marcadas (nomes repetidos), e que (ii) a RNP possa ser um princípio de processamento cognitivo universal, no sentido de independente do parâmetro *pro-drop*.

Por sua vez, o estudo das línguas românicas mostra-se potencialmente esclarecedor para tais fenômenos sob investigação, devido à sua maior riqueza flexional, que pode ser mobilizada durante o processo de resolução da referência.

É sabido que o chinês possui um sistema morfossintático bastante empobrecido, carente de marcas de concordância verbal, o que, conseqüentemente, obriga seus falantes a dependerem fortemente de pistas contextuais e pragmáticas para a interpretação de pronomes nulos, ao passo que o uso do pronome pleno é discursivamente motivado e justificado, uma vez que contém traços semânticos (como pessoa, gênero e número) que auxiliam na identificação de um referente.

Já nas línguas românicas, como o italiano, o espanhol e o português,⁴ as pistas semânticas oferecidas aos falantes pelos pronomes plenos são redundantes (i.e., não introduzem informação nova), visto que esses traços já estão codificados nas desinências verbais, ao contrário do que ocorre em chinês. Os pronomes nulos, por sua vez, apesar de não serem redundantes, também não adicionam informação nova sobre seus referentes, podendo seus antecedentes ser facilmente inferidos através das marcas de concordância verbal:

(...) as línguas que admitem uma categoria vazia em posição de sujeito de uma sentença com tempo finito têm também [geralmente] um paradigma verbal capaz de identificar essa categoria vazia de forma a lhe garantir a interpretação correta (FIGUEIREDO SILVA, 1996, p. 29).

Em decorrência disso, algumas questões que demandam investigação são: por que as gramáticas dessas línguas licenciam tanto a existência de pronomes plenos quanto de nulos? O que justifica a existência de ambas as possibilidades de retomada anafórica? Estariam elas também em variação livre nas línguas românicas?

Em resposta a alguns desses questionamentos, Alonso-Ovalle et al. (2002) e Gelormini-Lezama (2008) oferecem evidências de que, ao menos em espanhol, pronomes plenos e nulos não se encontram em variação livre (como em chinês), mas sim em distribuição complementar.

Em um estudo de questionário realizado por Alonso-Ovalle et al., solicitava-se a falantes nativos de espanhol que fossem lidas frases como as que se seguem (2002, p. 3-4):

- (1) “Juan pegó a Pedro. (pro) Está enfadado.”
- (2) “Juan pegó a Pedro. Él está enfadado.”

Em seguida, os informantes deveriam responder a uma pergunta sobre qual era o referente (Juan ou Pedro) do sujeito gramatical da segunda sentença. Quando esse sujeito era nulo, como em (01), a maioria das respostas (73,2%) apontou o sujeito da primeira sentença como o antecedente correto; por outro lado, quando esse sujeito era um pronome foneticamente pleno, como em (02), a porcentagem de respostas indicando o sujeito da primeira sentença como antecedente caiu para 50,2%. Os resultados dessa pesquisa, de mais três estudos de questionário e de um estudo envolvendo julgamentos mostraram que o espanhol exibe a mesma diferença que o italiano (CARMINATI *apud* ALONSO-OVALLE et al. 2002) no tocante ao comportamento de

⁴ Considerações acerca das mudanças pelas quais tem passado o PB, no tocante ao seu paradigma verbal, serão feitas posteriormente neste artigo.

expressões anafóricas reduzidas: nulos buscam seus antecedentes preferencialmente na posição sintática mais alta da oração (Spec, IP), ao passo que pronomes plenos tendem a estabelecer correferência com antecedentes em posições sintáticas inferiores.⁵

Todavia, questionários e tarefas de julgamento como os empregados pelos autores correspondem a métodos *off-line* de experimentação, isto é, que não são sensíveis ao processamento linguístico em tempo real e, portanto, não são comparáveis aos estudos sobre o inglês e o chinês reportados neste artigo. Além disso, não estava no escopo do trabalho de Alonso-Ovalle et al. mostrar se a RNP se manifestaria também em espanhol. Esta tarefa foi levada a cabo por Gelormini-Lezama (2008).

O autor elaborou dois experimentos de leitura autocadenciada com vistas a investigar, em espanhol, o efeito de RNP e a carga de processamento de nomes repetidos, pronomes plenos e nulos em retomadas anafóricas com antecedentes em posição de sujeito e de objeto. Segundo análise de Gelormini-Lezama, seus resultados comprovaram a existência de RNP, já que, na retomada de antecedentes em posição de sujeito, nomes repetidos foram processados mais lentamente em comparação com pronomes nulos, fato que fortalece o argumento de que essa penalidade possa ser, de fato, de natureza linguística universal.⁶

Curiosamente, os resultados também revelaram que, para tais antecedentes em posição de sujeito, são os pronomes plenos que elicitam uma penalidade de processamento, de natureza linguística específica (até agora, descoberta apenas em espanhol), que foi nomeada de *overt pronoun penalty* (OPP, ou penalidade do pronome pleno): elipses anafóricas são preferidas a pronomes plenos, sendo estes penalizados por não serem capazes de oferecer nenhuma vantagem diferencial em um contexto onde um antecedente pode ser identificado sistematicamente através de marcas de concordância verbal.

Em outras palavras, isso significa dizer que, em espanhol, passagens como (03) são processadas mais facilmente em relação a passagens como (04), que por sua vez são mais facilmente processadas em relação a (05), conforme exemplos extraídos de Gelormini-Lezama (2008, p. 17):

- (3) “Juan se encontró con María. La vio triste.”
- (4) “Juan se encontró con María. Juan la vio triste.”
- (5) “Juan se encontró con María. Él la vio triste.”

Já nos casos de correferência com antecedentes em posição de objeto, as diferenças nos modos como nomes repetidos e pronomes plenos foram processados em espanhol não foram estatisticamente significativas, tendo sido nulos as formas anafóricas mais penalizadas em comparação com as outras duas, elicitando os tempos de reação mais elevados dentre todas as condições testadas.

Português e espanhol são línguas bastante semelhantes morfológica e sintaticamente, o que levanta questionamentos sobre se o padrão encontrado para o espanhol também corresponderia ao padrão de processamento de expressões anafóricas existente em PB. Essa

⁵ Para Alonso-Ovalle et al. (2002), que se valem da *Position of Antecedent Hypothesis* ou PHA (CARMINATI *apud* ALONSO-OVALLE et al. 2002), a saliência de um antecedente é determinada exclusivamente pela posição sintática por ele ocupada, sendo propriedades morfológicas, semânticas, pragmáticas e cognitivas totalmente desconsideradas na explicação do fenômeno da correferência.

⁶ Cabe notar que Gelormini-Lezama (2008) interpreta seus dados à luz da *Informational Load Hypothesis* ou ILH (ALMOR, 1999), que propõe que o custo processual de uma anáfora precisa estar justificado em termos de sua função discursiva dentro de um contexto específico. A ILH também postula a existência de uma relação entre processamento linguístico e memória de trabalho verbal. Assim, processar um nome repetido implica processar mais material fonológico e reativar mais informação do que o necessário para recuperar um referente que já está em foco, desperdiçando, portanto, espaço precioso na memória. Por outro lado, formas referenciais menos informativas e mais leves do ponto de vista processual (como pronomes plenos e nulos) não só são capazes de garantir uma recuperação correta de um antecedente saliente como até mesmo o fazem de maneira mais eficiente do que nomes repetidos, economizando recursos dos sistemas de memória e atenção.

comparação torna-se ainda mais instigante porque o PB tem passado por uma rápida simplificação em sua morfologia verbal, o que estaria levando a um uso cada vez mais frequente do sujeito explícito e ao virtual desaparecimento do sujeito nulo pessoal, conforme tem sido vastamente atestado por diversos autores, dentre eles Duarte (In: ROBERTS; KATO, 1996, p. 107) e Tarallo (In: ROBERTS; KATO, 1996, p. 89), respectivamente:

De fato, a mudança que se observa no português do Brasil, que parece estar evoluindo de uma marcação positiva para uma marcação negativa dentro do parâmetro 'pro-drop', coincide com significativa redução ou simplificação nos paradigmas flexionais.

Com base nos resultados obtidos em Tarallo (1983, 1985), que atestavam o crescimento de sujeitos lexicais acompanhado de um decréscimo no objeto direto anafórico, um argumento forte pode ser feito em relação à modalidade brasileira como um sistema em fase de transição de língua 'pro-drop' para 'não pro-drop', isto é, uma mudança paramétrica.

Diante disso, é possível prever que, em PB, diferentemente do que foi atestado em espanhol, sujeitos nulos correferenciais sejam as formas anafóricas mais penalizadas (i.e., as mais difíceis de serem processadas), já que o antecedente não pode mais ser inferido através dos traços de concordância verbal. Seria também de se esperar que o processamento de nomes repetidos, supostamente regido por princípios universais, fosse penalizado no contexto em que seus antecedentes são sintaticamente salientes (i.e., realizam-se em posição de sujeito). Nesse contexto, esperar-se-ia que, em uma escala decrescente de facilidade de processamento, nulos fossem seguidos por nomes repetidos e, por fim, pronomes, que seriam, então, a forma anafórica mais favorecida cognitivamente; em outras palavras, esperar-se-ia verificar a existência de RNP, mas não de OPP. Já na retomada de antecedentes não salientes (i.e., em posição de objeto), a expectativa seria de que não houvesse diferenças significativas entre nomes repetidos e pronomes, dado que essas diferenças não foram verificadas nem em inglês, nem em chinês e nem em espanhol.

Até onde é do conhecimento dos autores deste artigo, os principais, dentre os poucos trabalhos psicolinguísticos desenvolvidos e publicados em PB sobre os problemas acima esboçados, são os de Leitão (2005) e Queiroz e Leitão (2008).

Em Leitão (2005), são reportados dois experimentos de leitura autocadenciada, dos quais o de maior interesse para nós é o primeiro, por ser o que busca mais diretamente investigar a existência ou não de RNP em PB. Nesse experimento, o autor investigou apenas o processamento de nomes repetidos e pronomes plenos (formas anafóricas nulas não foram alcançadas pelo escopo da pesquisa), sendo que tanto a expressão anafórica quanto seu antecedente se encontravam em posição de objeto direto. Segundo o autor, seus resultados indicaram a existência de RNP em PB, mostrando, ao contrário de Gordon, Grosz e Gilliom (1993), que nomes repetidos, em comparação com pronomes plenos, podem também ser penalizados em retomadas anafóricas em posição de objeto.

Entretanto, acreditamos que o trabalho de Leitão (2005), apesar de seu caráter seminal em PB, não possa ser diretamente comparado ao de Gordon, Grosz e Gilliom (1993), pois o fato de seu experimento ter mostrado a ocorrência de RNP em retomadas com antecedentes objetos nada revela sobre a existência dessa penalidade nos outros contextos sintáticos, como os que têm sido classicamente estudados desde Gordon, Grosz e Gilliom (1993), isto é, contextos de retomadas anafóricas com antecedentes em posição de sujeito.

No contexto sintático mencionado, nomes repetidos e pronomes plenos foram investigados pela primeira vez em PB por Queiroz e Leitão (2008), que conduziram outros dois experimentos de leitura autocadenciada, bastante semelhantes aos reportados em Leitão (2005). Os resultados do primeiro experimento mostraram que nomes repetidos foram penalizados em relação a pronomes plenos, elicitando tempos de reação significativamente maiores. Portanto, foi

verificada a existência de RNP em retomadas anafóricas também na posição de sujeito, e não só na de objeto, conforme já demonstrado apenas por Leitão (2005).

Entretanto, mesmo testando a posição de sujeito, acreditamos que o trabalho de Queiroz e Leitão (2008) também não permite que os resultados obtidos sejam diretamente comparados aos de Gordon, Grosz e Gilliom (1993), uma vez que os primeiros não investigaram a natureza do processamento correferencial de nomes repetidos e pronomes plenos fazendo uso das mesmas estruturas linguísticas utilizadas pelos últimos. Em primeiro lugar, os itens experimentais de Queiroz e Leitão (2008) consistiram em duas orações coordenadas, fragmentadas em nove segmentos (com o sexto segmento como crítico); todavia, quatro dos cinco experimentos descritos em Gordon, Grosz e Gilliom (1993) e os demais estudos envolvendo experimentos de leitura autocadenciada aqui reportados utilizaram orações independentes, não fragmentadas em segmentos menores, ou seja, tais estudos tiveram como estímulos orações inteiras, e não fragmentos de oração.⁷ Em segundo lugar, os seus itens experimentais continham anáforas intrassentenciais, isto é, com antecedentes dentro de uma mesma sentença; no entanto, são anáforas intersentenciais (com antecedentes presentes em outra sentença) que têm sido o tipo de forma referencial estudado nos trabalhos reportados sobre o inglês, o chinês e o espanhol.

Diante do que foi exposto, com vistas a (i) investigar o comportamento de nomes repetidos, pronomes plenos e pronomes nulos, (ii) verificar a existência ou não de RNP e OPP e, por fim, (iii) oferecer uma análise do fenômeno da correferência em PB que possa ser diretamente comparada aos trabalhos de Gordon, Grosz e Gilliom (1993), Yang et al. (1999) e Gelormini-Lezama (2008), elaboramos um experimento de leitura autocadenciada, cujo delineamento⁸ corresponde ao do primeiro experimento descrito em Yang et al. (1999).

Um experimento no paradigma experimental de leitura autocadenciada (*self-paced reading*, ou ainda leitura automonitorada) consiste na medida dos tempos de reação (*reaction times* ou RTs) de sujeitos na leitura seja de palavras, seja de fragmentos ou sintagmas, ou ainda, como foi o caso deste experimento, de sentenças inteiras de um dado conjunto de textos. Esse tempo de reação corresponde à medida do intervalo entre a apresentação de um estímulo linguístico na tela do computador, sua leitura e o acionamento posterior de uma tecla. Além disso, o tempo de exposição de cada sentença é controlado pelo próprio sujeito, sendo medido com grande precisão, na ordem de milésimos de segundo (ms). A premissa subjacente a esse paradigma é a de que tempos de reação baixos são indicativos de facilidade no processamento de uma determinada estrutura linguística, ao passo que quanto maiores os tempos de reação, assume-se que mais difícil foi o processamento.

Um exemplo das sentenças utilizadas como estímulos linguísticos no experimento que conduzimos encontra-se na Tabela 1. Como se vê, esses estímulos consistiram em passagens compostas por três sentenças, constituintes de um minidiscorso semanticamente coerente, sendo a segunda sentença o local onde se deram as manipulações das variáveis independentes, a saber, forma anafórica e posição sintática. Foram atribuídos diferentes tipos de expressões referenciais (nomes repetidos, pronomes e nulos) a diferentes posições sintáticas (sujeito e objeto), tendo diferentes combinações entre os níveis das duas variáveis independentes gerado um delineamento com seis condições experimentais, a saber: Nome-Nome, Pronome-Nome, Nulo-Nome, Nome-Pronome, Pronome-Pronome, e Nulo-Pronome. O tempo de reação (doravante RT) da segunda sentença de cada passagem foi a variável dependente do experimento, isto é, a variável

⁷ A necessidade de se levar em consideração essas diferenças no modo de apresentação dos estímulos linguísticos encontra respaldo, por exemplo, no estudo de Nair e Almor (*apud* GELORMINI-LEZAMA, 2008), que forneceu evidências de que a RNP ocorre em função de processos integrativos (e não locais) de leitura, não sendo, portanto, necessariamente elicitada no segmento ou sintagma em que a forma anafórica se localiza.

⁸ Por *delineamento*, entenda-se *delineamento experimental*, termo que se refere ao plano para se atribuir sujeitos a condições experimentais e à análise estatística associada a esse plano. Um desenho experimental também especifica as variáveis dependentes e independentes, bem como os ruídos ou vieses, ou seja, as fontes que influenciam negativamente a variável dependente e que ameaçam a validade das generalizações feitas a partir dos dados. Para uma boa introdução ao tema, ver Kirk (1995).

mensurável, onde se esperava encontrar alterações em função das manipulações nas variáveis independentes.

TABELA 1
Exemplo de estímulo experimental

SENTENÇA INICIAL:	
Bruna namora Vítor há mais de dois anos e meio.	
CONDIÇÕES:	SENTENÇA CRÍTICA (6 VERSÕES):
Nome-Nome	Bruna conheceu Vítor em uma viagem à França.
Pronome-Nome	Ela conheceu Vítor em uma viagem à França.
Nulo-Nome	Conheceu Vítor em uma viagem à França.
Nome-Pronome	Bruna conheceu-o em uma viagem à França.
Pronome-Pronome	Ela conheceu-o em uma viagem à França.
Nulo-Pronome	Conheceu-o em uma viagem à França.
SENTENÇA FINAL:	
É possível encontrar um grande amor em outro país.	

1. O experimento

1.1 Participantes

62 voluntários falantes nativos de PB participaram do experimento, dentre os quais 36 do sexo feminino e 24 do sexo masculino. A idade média dos sujeitos foi de 25 anos, variando entre 17 e 46 anos.

1.2 Estímulos

Foi elaborado um conjunto de 42 passagens experimentais, como a que está exemplificada na Tabela 1. Conforme mencionado anteriormente, cada passagem continha três sentenças, constituindo um minidiscorso semanticamente coerente e obedecendo aos seguintes controles:

- A primeira sentença introduzia, através de nomes próprios dissílabos, dois indivíduos de gêneros distintos, dispostos na ordem sujeito-verbo-objeto (SVO), de modo que uma entidade se realizava na posição de sujeito, enquanto a outra, na de objeto (direto ou indireto). A escolha de nomes próprios de gêneros diferentes e da ordem SVO se deu pautada, respectivamente, na necessidade de se eliminar qualquer influência de ambiguidade de gênero sobre o processo de resolução anafórica e no fato de a ordem SVO ser, atualmente, a ordem sintática não marcada em PB. Além disso, metade das passagens experimentais continha nomes próprios masculinos seguidos por femininos, enquanto na outra metade essa sequência se invertia, ou seja, nomes femininos seguidos por masculinos.

- Havia uma relação de paralelismo sintático entre as duas primeiras sentenças, isto é, as funções sintáticas exercidas pelas entidades na sentença inicial eram mantidas na segunda sentença, a sentença crítica, da qual foram construídas seis diferentes versões, a partir de associações entre os níveis das variáveis independentes. Além disso, as sentenças críticas tiveram seus tamanhos controlados em função do número de sílabas. Cumpre também mencionar que pronomes nulos não foram utilizados em posição de objeto para que o desenho experimental reproduzisse fielmente aquele utilizado por Yang et al. (1999) para o chinês, correspondendo este, por sua vez, exatamente ao desenho utilizado por Gordon, Grosz e Gilliom (1993), com exceção

da inclusão de formas nulas na condição sujeito. Desse modo, os resultados deste experimento serão diretamente comparáveis aos resultados já obtidos em chinês e inglês.

▪ A terceira sentença, de conteúdo genérico, não fazia menção a nenhuma das duas entidades presentes nas sentenças anteriores, tendo sido incluída de modo que a sentença crítica não aparecesse no fim da passagem, local onde o RT é naturalmente mais longo devido a processos semântico-discursivos de integração.

Foi também construído um conjunto adicional de 84 passagens distratoras (ou *fillers*), destinado a mascarar as passagens experimentais, disfarçando, para os sujeitos, qual o objetivo do teste. Assim como as passagens experimentais, as distratoras também eram compostas por três sentenças, contendo referentes introduzidos por nomes próprios ou outros tipos de descrições definidas e indefinidas, constituindo minidiscursos semanticamente coerentes. Por fim, cabe notar que tais passagens distratoras possuíam estruturas diversas, não diretamente relacionadas às manipulações experimentais.

1.3 Procedimento

O experimento foi rodado em um computador pessoal, em uma única sessão de aproximadamente quarenta minutos de duração. Os sujeitos foram instruídos a ler cada sentença no seu ritmo mais natural possível, nem muito apressada, nem muito vagarosamente. Após lerem as instruções na tela, eles iniciavam uma sessão de treinamento com cinco passagens distratoras, para que pudessem se familiarizar com a tarefa de leitura autocadenciada, apreender a parte mecânica da tarefa antes do início do experimento propriamente dito e também tirar suas dúvidas.

Ao término da sessão de treinamento, os sujeitos eram avisados de que, ao pressionarem a barra de espaço, o experimento teria início. Antes da apresentação de cada passagem (tanto experimental quanto distratora), aparecia a seguinte sentença no centro da tela do computador: “Pressione a barra de espaço para ler a oração seguinte”. Ao pressionarem a barra de espaço, os sujeitos viam então a primeira sentença da passagem. Ao pressionarem novamente a barra de espaço quando do término da leitura, a sentença anterior desaparecia da tela e era substituída, no mesmo local, pela sentença crítica (sendo o tempo de reação desta sentença a variável dependente do experimento), e assim sucessivamente, até que os sujeitos chegassem à terceira sentença.

Ao final de cada passagem, isto é, após a leitura da terceira sentença, uma pergunta de compreensão aparecia no centro da tela. Os sujeitos respondiam a ela pressionando ou a barra de espaço, para dizerem “sim”, ou a tecla *shift* (direita ou esquerda, a depender de o sujeito ser destro ou canhoto), para responderem “não”. Essas perguntas de compreensão, incluídas também nos estímulos de treinamento, não exigiam a rememoração dos nomes próprios mencionados, mas sim das estruturas dos eventos descritos nas passagens. Além disso, elas foram elaboradas de modo que metade das respostas corretas fosse “sim” e a outra metade fosse “não”. Suas funções eram, por um lado, contribuir para disfarçar o objetivo do experimento e, por outro, garantir a atenção dos sujeitos durante toda a tarefa de leitura.

Foi utilizado o programa *E-prime*⁹ para aleatorizar, diferentemente para cada sujeito, a ordem dos estímulos experimentais e distratores a ser apresentada na tela do computador, de modo a eliminar possíveis efeitos da ordem de apresentação dos estímulos nos resultados. O programa também fazia com que as sentenças críticas das passagens experimentais fossem apresentadas a cada sujeito em apenas uma das seis condições possíveis; ou seja, em cada passagem, os sujeitos só viam uma dentre as seis possíveis versões da sentença crítica. No entanto, considerando o conjunto de todos os sujeitos, todas as sentenças críticas das passagens ocorreram igualmente em todas as seis condições testadas.

⁹ Maiores informações sobre o programa estão disponíveis em: <<http://www.pstnet.com/eprime.cfm>> (acesso em: 9 abr. 2012).

1.4 Resultados

A Tabela 2 (região central) mostra as médias dos RTs da segunda sentença em todas as condições testadas (Nome-Nome, Pronome-Nome, Nulo-Nome, Nome-Pronome, Pronome-Pronome e Nulo-Pronome).¹⁰ A partir desses dados, foram calculadas, para cada uma das formas referenciais (nome repetido, pronome pleno e nulo) associadas a antecedentes na posição de sujeito, as médias dos RTs das sentenças críticas que as continham (independentemente da forma referencial presente na posição de objeto dessas sentenças). Do mesmo modo, foram calculadas as médias dos RTs das sentenças críticas que continham nomes repetidos e pronomes plenos com antecedentes na posição de objeto (independentemente da forma referencial presente na posição de sujeito). As médias calculadas através desse segundo procedimento podem ser visualizadas tanto na Tabela 2 (região marginal) quanto na Figura 1.

Os RTs das sentenças críticas contendo nomes repetidos e pronomes plenos como sujeitos gramaticais foram, respectivamente, 2674 e 2659 ms, sendo os primeiros, portanto, apenas 15 ms mais lentos em comparação com os segundos. Por sua vez, o processamento das sentenças contendo sujeitos nulos foi comparativamente mais rápido, com diferenças superiores a 250 ms em relação a nomes repetidos e pronomes plenos. Já as sentenças contendo nomes repetidos e pronomes plenos como objetos gramaticais levaram, respectivamente, 2641 e 2501 ms para serem processadas, tendo sido a diferença entre os RTs de apenas 140 ms.

TABELA 2
Médias dos RTs (em ms) das sentenças críticas segundo forma anafórica e posição sintática

POSIÇÃO SINTÁTICA DO ANTECEDENTE		OBJETO		MÉDIA
		<i>Nome repetido</i>	<i>Pronome</i>	
SUJEITO	<i>Nome repetido</i>	2737 ± 109	2612 ± 114	2674 ± 79
	<i>Pronome</i>	2724 ± 113	2594 ± 115	2659 ± 81
	<i>Nulo</i>	2463 ± 101	2298 ± 115	2381 ± 77
MÉDIA		2641 ± 63	2501 ± 67	

¹⁰ Do conjunto inicial de 62 sujeitos, 2 foram excluídos, devido aos seus índices de acerto nas respostas às perguntas de compreensão terem sido inferiores a 80%, limite de corte estabelecido como aceitável a partir da própria média de acertos do conjunto de todos os informantes (=87%, variando de 80 a 98%).

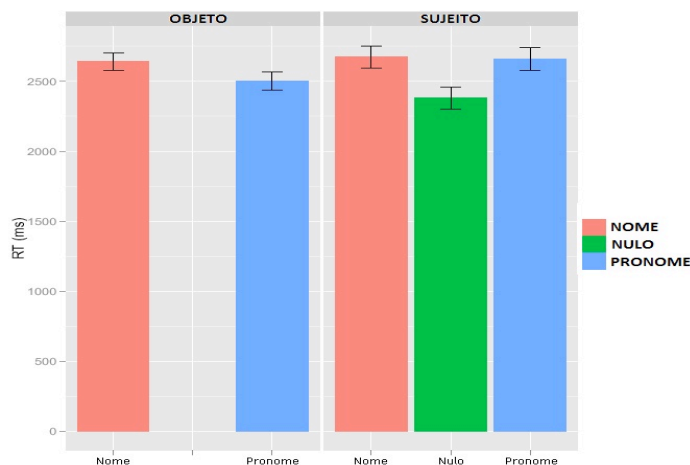


FIGURA 1 – Médias dos RTs (em ms) das sentenças críticas segundo posição sintática

De modo a testar os graus de significância de tais diferenças, procedemos a uma análise estatística de variância (ANOVA) dos dados, conduzida tanto por sujeitos (F1) quanto por itens experimentais (F2). Os resultados revelaram que, quando a expressão referencial retomava antecedentes sintaticamente salientes, o efeito de forma anafórica foi significativo tanto por sujeitos (F1 (2,118) = 14.8, $p < 0.001$) quanto por itens (F2 (2,82) = 10.1, $p < 0.001$). Similarmente, estando as expressões referenciais associadas a antecedentes sintaticamente não salientes, o efeito de forma anafórica atingiu significância por sujeitos (F1 (1,59) = 11.1, $p < 0.01$) e por itens (F2 (1,41) = 6.1, $p < 0.05$). Por fim, a interação de forma anafórica e posição sintática do antecedente não se mostrou significativa nem por sujeitos (F1 (2,118) = 0.09, $p = 0.92$), nem por itens (F2 (2,82) = 0.1, $p = 0.90$).

Em seguida, com vistas a explorar mais detalhadamente as interações entre os níveis de posição sintática (sujeito e objeto) e de forma anafórica (nomes repetidos, pronomes plenos e nulos), procedemos a um teste *post hoc* de Fisher. Os resultados mostraram que (i) nomes repetidos e pronomes plenos apresentaram comportamentos idênticos entre si, não havendo as manipulações do fator posição sintática gerado efeitos significativos sobre eles ($p > 0.05$), e que (ii) só houve diferença significativa entre nulos sujeitos e demais condições ($p < 0.001$).

Em síntese, os resultados revelaram que pronomes nulos que se realizam na posição de sujeito e têm antecedentes que também são sujeitos gramaticais de suas respectivas sentenças são as formas mais favorecidas do ponto de vista do processamento correferencial, ao passo que nomes repetidos e pronomes plenos, independentemente da saliência sintática de seus antecedentes, são igualmente penalizados entre si.

2. Discussão

As hipóteses iniciais não foram inteiramente corroboradas pelos dados. Devido à simplificação do paradigma flexional verbal pela qual o PB tem passado, a identificação dos sujeitos nulos não pode mais se dar através das desinências de concordância, o que conduziu à previsão de que pronomes nulos seriam as formas correferenciais mais penalizadas. Esperava-se também encontrar RNP, devido à alegação (YANG et al., 1999; GELORMINI-LEZAMA, 2008) de que essa penalidade se trataria, provavelmente, de um universal linguístico, visto que foi verificada tanto em línguas totalmente *pro-drop* (como o chinês e o espanhol), quanto em línguas não *pro-drop* (como o inglês). Por conseguinte, esperava-se que pronomes plenos associados a antecedentes sintaticamente salientes fossem as formas anafóricas processadas mais rapidamente.

Todavia, em retomadas anafóricas na posição de sujeito, sujeitos nulos foram as formas mais fáceis de serem processadas pelos leitores, com nomes repetidos e pronomes plenos sendo

igualmente penalizados entre si, o que indica, de acordo com nossa análise, a existência de OPP, mas não de RNP. Esse resultado se contrapõe aos de Gordon, Grosz e Gilliom (1993) e aos de Queiroz e Leitão (2008), pois que foi verificado que a RNP não possui natureza universal, não se estendendo ao PB, e também aos de Yang et al. (1999), já que encontramos diferenças significativas entre pronomes nulos e pronomes plenos. O padrão de processamento encontrado guarda, contudo, semelhanças com os resultados provenientes do espanhol, sistema linguístico em que, de acordo com Gelormini-Lezama (2008), pronomes nulos são processados mais facilmente do que nomes repetidos, que por sua vez são processados mais facilmente do que pronomes plenos (quando da retomada de antecedentes sintaticamente salientes). Portanto, tanto em PB quanto em espanhol, pronomes nulos foram penalizados em comparação com nomes repetidos.

Gelormini-Lezama (2008) interpreta essa penalização como uma instância de RNP. Nós, contudo, não a interpretamos da mesma maneira, e tal análise decorre da opção teórica de nos mantermos fiéis à caracterização original da RNP feita por Gordon, Grosz e Gilliom (1993). Nesse estudo clássico, verificou-se que nomes repetidos foram penalizados em relação a pronomes plenos (visto que o inglês não autoriza a existência de pronomes nulos, salvo nos contextos supostamente universais), e foi a partir de tal dado que a RNP foi nomeada e caracterizada. Mesmo em estudos posteriores que incluíram a categoria dos pronomes nulos no seu escopo de investigação, como o de Yang et al. (1999), a penalização de nomes repetidos foi verificada em relação à classe de expressões referenciais reduzidas (pronomes plenos e nulos) como um todo, e não em relação a membros isolados da classe, como o foi em Gelormini-Lezama (2008).

Ainda que o modo de interpretação da RNP não seja consensual, a preferência pelo nulo verificada em espanhol e em PB pode ser explicada dentro do quadro teórico proposto por Gundel, Hedberg e Zacharski (1993): quanto mais saliente for o antecedente, \square sendo a saliência, no caso do nosso experimento, determinada pela posição sintática do antecedente \square menos informativo precisa ser o termo anafórico; em outras palavras, nulos são as formas mais eficientes na retomada de referentes que estão em foco, no centro atencional dos falantes. A explicação para a preferência pelo nulo pode também ser buscada na relação entre processamento linguístico e memória de trabalho verbal proposta por Almor (1999): processar um nome repetido ou um pronome implica processar mais informação do que o necessário para recuperar um referente saliente; nesse contexto, uma forma referencial nula, bem mais leve do ponto de vista do processamento, é capaz de garantir a recuperação correta do antecedente, sendo o seu uso justificado, portanto, em termos de economia de recursos dos sistemas de atenção e memória de trabalho verbal.

Já em retomadas anafóricas na posição de objeto, não verificamos diferenças significativas nos comportamentos de nomes repetidos e pronomes plenos, dado que contrasta com os achados de Leitão (2005), mas que vai de encontro às hipóteses de trabalho e que se aproxima dos resultados de Gordon, Grosz e Gilliom (1993), de Yang et al. (1999) e de Gelormini-Lezama (2008).

Analisando os padrões de processamento que foram encontrados, relacionados à retomada de antecedentes tanto sintaticamente salientes quanto não salientes, levantamos uma objeção importante: poderiam ter as diferenças inerentes ao tamanho das sentenças críticas (em suas seis possíveis versões) enviesado os dados, ou seja, poderiam elas ter interferido negativamente na manipulação da variável dependente? No experimento que conduzimos, bem

como no de Yang et al. (1999), o número total de sílabas das sentenças críticas variou em função das manipulações de forma anafórica (rever Tabela 1). Desse modo, houve sentenças com 16 sílabas nas condições Nome-Nome e Pronome-Nome; com 14 sílabas na condição Nulo-Nome; com 15 sílabas nas condições Nome-Pronome e Pronome-Pronome; e, por fim, sentenças com 13 sílabas na condição Nulo-Pronome. Tais diferenças foram inevitáveis, porquanto ocorreram em função de diferenças no tamanho dos nomes repetidos empregados (todos dissílabos) e dos pronomes tônicos (também dissílabos) em relação aos pronomes clíticos (monossilábicos) e aos pronomes nulos (elipses).

Todavia, descartamos tal objeção levantada, visto que a literatura fornece inúmeras evidências (YANG et al., 1999; GELORMINI-LEZAMA, 2008, *inter alia*) de que pequenas diferenças no número de sílabas das sentenças não são capazes de interferir na magnitude das penalidades de processamento reportadas: por um lado, em inglês (GORDON; GROSZ; GILLIOM, 1993) e em chinês (YANG et al., 1999), nomes repetidos foram penalizados em relação a pronomes plenos apenas quando da retomada de antecedentes em posição de sujeito, e não em posição de objeto, mas as manipulações de forma anafórica geraram, contudo, as mesmas diferenças de tamanho em ambas as posições sintáticas envolvidas, nos dois experimentos; e, por outro lado, nos dados provenientes do espanhol (GELORMINI-LEZAMA, 2008), sentenças contendo formas referenciais nulas foram as processadas mais rapidamente na retomada de antecedentes em posição de sujeito, mas, ao mesmo tempo, as mais penalizadas na retomada de antecedentes em posição de objeto, não tendo havido nenhuma diferença no tamanho das sentenças contendo pronomes nulos em ambas as condições de retomada que explicasse tal inversão na escala de preferência processual.

Em síntese, nossos resultados colocam o PB em uma posição singular frente às demais línguas que têm sido objeto de estudo em psicolinguística comparada: o padrão de processamento das expressões anafóricas estudadas não se assemelha nem ao inglês, nem ao chinês, nem completamente ao espanhol. No entanto, é com este último sistema linguístico que o PB apresentou mais similaridades, muito provavelmente em função da proximidade lexical, morfológica e sintática que as duas línguas apresentam, advinda de sua origem comum.

Os resultados acabam também por ultrapassar o domínio do processamento correferencial e levantam questionamentos sobre se haveria ocorrido, de fato, uma mudança dentro do parâmetro *pro-drop* em PB (ao menos para os tipos de passagens estudadas neste experimento), a despeito do que vem sendo vastamente afirmado na literatura linguística. Nesse ponto, é necessário que a pesquisa psicolinguística dialogue com estudos sociolinguísticos sobre o tema, de modo que se possam encontrar corroborações empíricas para as evidências experimentais obtidas.

Um desses estudos é o de Cabana (2004),¹¹ que investigou o comportamento da variável sujeito pronominal nulo ~ sujeito pronominal pleno na variedade de PB falada em Belo Horizonte (MG). A autora realizou três estudos variacionistas, dois em tempo aparente e um em tempo real, chegando à conclusão de que há um equilíbrio nos usos das variantes, com um leve crescimento no uso do sujeito nulo de 1980 a 2004, sendo essa preferência maior nas gerações mais jovens, em ambos os momentos.

¹¹ Além de Cabana (2004), outro trabalho sociolinguístico que também apresenta evidências empíricas contra a ocorrência de uma mudança paramétrica em progresso é Nicolau (1995).

No seu *corpus* de 2004, obtido por meio de entrevistas sociolinguísticas com falantes belo-horizontinos, foram encontrados 55% de sujeitos nulos e 45% de sujeitos pronominais plenos, indicando, portanto, equilíbrio no comportamento das variantes, com leve preferência pelos sujeitos nulos. Logo, não se identificou o abandono de uma forma mais arcaica (supostamente o sujeito nulo) em favor de outra inovadora (sujeito pronominal pleno). Já no *corpus* da década de 1980, também constituído por dados de fala, obtidos através de entrevistas sociolinguísticas com falantes belo-horizontinos no período de 1982 a 1984, foram encontrados índices de 49% de sujeitos nulos e de 51% de sujeitos pronominais plenos; ou seja, as variantes também se encontravam, há trinta anos, em equilíbrio na comunidade de fala de Belo Horizonte. Procedendo-se a uma comparação entre os dois *corpora*, é perceptível um ligeiro crescimento no uso de sujeitos nulos de 1980 para 2004, e uma queda no uso de sujeitos pronominais plenos, o que constitui evidência contra o desaparecimento do sujeito nulo no PB, ao menos em sua variedade belo-horizontina.¹²

Além disso, os fatores linguísticos selecionados pelo programa VARBRUL como condicionadores do comportamento da variável em questão, nos dois *corpora* analisados, foram pessoa gramatical, tipo de oração e presença/ausência de elemento que antecede a posição de sujeito. No grupo de fatores sociais, apenas idade foi selecionado como relevante.¹³ De acordo com os resultados:

- As terceiras pessoas (tanto do singular quanto do plural) são altamente favorecedoras do sujeito nulo, com as maiores porcentagens de ocorrência de formas nulas associadas aos pesos relativos mais elevados.¹⁴

- Os tipos de oração mais favorecedores da ocorrência de sujeitos nulos são orações coordenadas não iniciais e orações independentes.

- A ausência de elementos antes da posição de sujeito favorece fortemente o uso do sujeito nulo, ao passo que a presença de algum elemento nessa posição o desfavorece, uma vez que dificulta a acessibilidade a um referente previamente introduzido.

- Quanto ao fator idade, foi encontrada uma preferência pelo uso do sujeito nulo na geração mais jovem (14 a 19 anos), tornando-se essa preferência cada vez menor ao longo das gerações intermediária (36 a 45 anos) e velha (55 a 67 anos), onde foi encontrado o menor índice de sujeitos nulos. Portanto, não se pode afirmar que a suposta variante inovadora (sujeito pronominal pleno) esteja mais presente na fala da geração mais jovem, evidência necessária, contudo, para se caracterizar, no tempo aparente, uma variação como constituindo mudança em progresso.

É interessante observar que, fortuitamente, os contextos linguísticos anteriormente descritos, demonstrados por Cabana (2004) como os mais favorecedores da ocorrência de sujeitos nulos, correspondem exatamente às características estruturais das passagens experimentais utilizadas no presente experimento de leitura autocadenciada. Esse fato pode

¹² É também relevante considerar que a maioria dos trabalhos sociolinguísticos que aponta para uma mudança em direção ao preenchimento da posição de sujeito apresenta dados provenientes de São Paulo (cf. TARALLO In: ROBERTS; KATO, 1996) e Rio de Janeiro (cf. DUARTE In: ROBERTS; KATO, 1996), ao contrário de Cabana (2004), cujos dados (assim como os do experimento reportado neste artigo) referem-se à variedade de PB falada em Belo Horizonte (MG). Logo, não se pode descartar a possibilidade de que as variantes sujeito nulo e sujeito pleno constituam um caso de variação diatópica, não atingindo homogeneamente todo o diassistema do PB.

¹³ Cabana (2004) realizou análises quantitativas posteriores incluindo o fator marcação verbal. Curiosamente, a despeito da suposta relação existente entre possibilidade de omissão do sujeito em sentenças finitas e marcas de concordância verbal, defendida, por exemplo, por Duarte (In: ROBERTS; KATO, 1996), esse fator não foi selecionado pelo programa VARBRUL como relevante para a variável sob investigação, não influenciando na escolha entre forma nula e forma plena.

¹⁴ Mesmo no trabalho de Duarte (In: ROBERTS; KATO, 1996), cujos dados, provenientes da comunidade de fala do Rio de Janeiro, apontam para um aumento no uso de sujeitos pronominais plenos, a terceira pessoa do singular apresentou índices bastante elevados de sujeitos nulos, em comparação com as demais pessoas do discurso.

explicar, a partir de uma perspectiva outra que não as das teorias psicolinguísticas e da referenciação, o porquê de sentenças contendo sujeitos nulos terem sido as mais favorecidas do ponto de vista do processamento correferencial em PB, ao contrário das hipóteses iniciais.

3. Considerações finais

A confirmação dos resultados experimentais reportados neste artigo exigirá a elaboração de novos experimentos (com diferentes delineamentos), bem como um levantamento mais aprofundado de estudos sociolinguísticos sobre o tema. Apesar disso, este estudo já lança luzes sobre questões cruciais do domínio do processamento mental da referência, domínio este que ainda se encontra relativamente subdescrito em PB.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALMOR, Amit. Noun-phrase anaphora and focus: the Informational Load Hypothesis. *Psychological Review*, v. 106, n. 4, p. 748-765, 1999.
- ALONSO-OVALLE, Luis et al. Null vs. overt pronouns and the topic-focus articulation in Spanish. *Journal of Italian Linguistics*, v. 14, n. 2, p. 151-169, 2002.
- CABANA, Nasle Maria. *Da realização do sujeito no Português do Brasil: um estudo em tempo real do uso do sujeito nulo na fala de Belo Horizonte/MG*. 2004. 77 f. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2004.
- CARMINATI, M. N. *The processing of Italian subject pronouns*. 2002. Tese (Doutorado em Linguística) – University of Massachusetts at Amherst, Amherst, 2002 *apud* ALONSO-OVALLE, Luis et al. Null vs. overt pronouns and the topic-focus articulation in Spanish. *Journal of Italian Linguistics*, v. 14, n. 2, p. 151-169, 2002.
- DUARTE, Maria Eugênia Lamoglia. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, Ian; KATO, Mary Aizawa (Org.). *Português Brasileiro: uma viagem diacrônica*. Campinas: Editora da UNICAMP, 1996.
- FIGUEIREDO SILVA, Maria Cristina. *A posição sujeito no Português Brasileiro: frases finitas e infinitivas*. Campinas: Editora da UNICAMP, 1996.
- GELORMINI-LEZAMA, Carlos. *Processing repeated names, overt pronouns and null reference in Spanish*. 2008. 46 f. Dissertação (Mestrado em Linguística) – College of Arts and Sciences, University of South Carolina, Columbia, 2008.
- GERNSBACHER, Morton Ann. Mechanisms that improve referential access. *Cognition*, v. 32, n. 2, p. 99-156, 1989.
- GORDON, Peter C.; GROSZ, Barbara J.; GILLIOM, Laura A. Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, v. 17, n. 3, p. 311-347, 1993.
- GUNDEL, Jeanette Kohn; HEDBERG, Nancy; ZACHARSKI, Ron. Cognitive status and the form of referring expressions in discourse. *Language*, v. 69, n. 2, p. 274-307, 1993.
- KIRK, Roger E. *Experimental design: procedures for the behavioral sciences*. Pacific Grove/CA: Brooks/Cole, 1995.
- LEITÃO, Márcio Martins. Processamento correferencial de nomes e pronomes em Português Brasileiro. *Revista Linguística (PPGL/UFRJ)*, v. 1, n. 2, p. 235-258, 2005.
- NAIR, Veena; ALMOR, Amit. On the measurement of the repeated name penalty with reading times and ERP. Não publicado *apud* GELORMINI-LEZAMA, Carlos. *Processing repeated names, overt pronouns and null reference in Spanish*. 2008. 46 f. Dissertação (Mestrado em Linguística) – College of Arts and Sciences, University of South Carolina, Columbia, 2008.
- NICOLAU, Eunice. *As propriedades de sujeito nulo e ordem V-S no Português Brasileiro*. 1995. 286 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1995.

- PRINCE, Ellen F. Toward a taxonomy of given-new information. In: COLE, Peter. *Radical Pragmatics*. New York: Academic Press, 1981.
- QUEIROZ, Karla Lima de; LEITÃO, Márcio Martins. Processamento do sujeito anafórico em Português Brasileiro. *Veredas – Revista de Estudos Linguísticos Online*, v. 2, n. 1, p. 163-166, 2008.
- TARALLO, Fernando. Diagnosticando uma gramática brasileira: o português d'aguém e d'além-mar ao final do século XIX. In: ROBERTS, Ian; KATO, Mary Aizawa (Org.). *Português Brasileiro: uma viagem diacrônica*. Campinas: Editora da UNICAMP, 1996.
- YANG, Chin Lung et al. Comprehension of referring expressions in Chinese. *Language and Cognitive Processes*, v. 14, n. 5/6, p. 715-743, 1999.